

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 11:51:58 ; Search time 50.08 seconds

(without alignments)  
75.940 Million cell updates/sec

Title: US-09-432-546-6  
Perfect score: 183  
Sequence: 1 RRPWWPKWPLIGGYPADPPPPP 26

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_17:\*  
1: sp.\_archaea:\*  
2: sp.\_bacteria:\*  
3: sp.\_fungi:\*  
4: sp.\_human:\*  
5: sp.\_invertebrate:\*  
6: sp.\_mammal:\*  
7: sp.\_mhc:\*  
8: sp.\_organelle:\*  
9: sp.\_phage:\*  
10: sp.\_plant:\*  
11: sp.\_rodent:\*  
12: sp.\_virus:\*  
13: sp.\_vertebrate:\*  
14: sp.\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78.5	42.9	1896	12 Q9DRA1	Q9DRA1 botrytis vi
2	77	42.1	161	10 Q08198	Q08198 nicotiana t
3	72	39.3	239	2 066118	066118 zymomonas m
4	70	38.3	172	5 Q23248	Q23248 caenorhabd
5	69	37.7	520	11 Q61078	Q61078 mus musculu
6	68.5	37.4	221	10 Q45946	Q45946 solanum tub
7	68	37.2	109	10 Q40548	Q40548 nicotiana t
8	68	37.2	142	3 Q9P3K2	Q9P3K2 neurospora
9	68	37.2	150	10 Q9FYZ6	Q9FYZ6 nicotiana t
10	68	37.2	196	10 Q08195	Q08195 nicotiana t
11	68	37.2	209	10 Q08194	Q08194 nicotiana t
12	68	37.2	615	10 Q9ZWK1	Q9ZWK1 nicotiana t
13	68	37.2	969	5 Q17795	Q17795 caenorhabd
14	67.5	36.9	352	5 Q94901	Q94901 dirosophila
15	66.5	36.3	157	10 Q08197	Q08197 nicotiana t
16	65	35.5	399	5 Q9GSG9	Q9GSG9 dictyostell
17	65	35.5	1662	2 P71431	P71431 leptothrix
18	64	35.0	148	11 Q64371	Q64371 rattus norv
19	64	35.0	650	10 Q49570	Q49570 arabidopsis

20	64	35.0	1137	10 Q9LNT8	Q9LNT8 arabidopsis
21	64	35.0	1420	2 Q52666	Q52666 escherichia
22	63.5	34.7	104	5 Q9VOW4	Q9VOW4 drosophila
23	63	34.4	172	5 Q00487	Q00487 hydra atten
24	63	34.4	1004	10 Q9L174	Q9L174 arabidopsis
25	62	33.9	92	10 Q24140	Q24140 nicotiana t
26	62	33.9	137	5 Q20468	Q20468 caenorhabd
27	62	33.9	137	10 Q9XG52	Q9XG52 lycopersico
28	62	33.9	390	4 Q9HAI3	Q9HAI3 homo sapien
29	62	33.9	502	4 Q9UG75	Q9UG75 homo sapien
30	62	33.9	584	10 Q23114	Q23114 arabidopsis
31	62	33.9	584	10 Q9SLH0	Q9SLH0 arabidopsis
32	62	33.9	738	3 P78943	P78943 schizosach
33	62	33.9	785	3 Q9HGP2	Q9HGP2 schizosach
34	62	33.9	856	4 Q76022	Q76022 homo sapien
35	62	33.9	856	4 Q9BUJ2	Q9BUJ2 homo sapien
36	61.5	33.6	671	5 Q9VBR3	Q9VBR3 drosophila
37	61.5	33.6	684	5 Q24035	Q24035 drosophila
38	61.5	33.6	688	10 Q9AYK0	Q9AYK0 oryza sativ
39	61	33.3	112	12 Q04273	Q04273 chimpanzee
40	61	33.3	112	12 Q76623	Q76623 human immun
41	61	33.3	112	12 P88144	P88144 human immun
42	61	33.3	112	12 Q88050	Q88050 chimpanzee
43	61	33.3	112	12 Q88057	Q88057 chimpanzee
44	61	33.3	112	12 Q88058	Q88058 chimpanzee
45	61	33.3	112	12 Q89552	Q89552 chimpanzee

## ALIGNMENTS

RESULT 1	
Q9DRA1	PRELIMINARY; PRT; 1896 AA.
AC Q9DRA1;	
DT 01-MAR-2001 (TREMBLrel. 16, Created)	
DR 01-MAR-2001 (TREMBLrel. 16, Last sequence update)	
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	
DE REPLICASE.	
OS Botrytis virus F.	
OC Viruses; ssRNA positive-strand viruses, no DNA stage.	
OX NCBI_TaxID:129395;	
RN [1]	
RP SEQUENCE FROM N.A.	
RX PubMed-1125160;	
RA Howitt R.L.J., Beaver R.E., Pearson M.N., Forster R.L.S.;	
RT "Genome characterization of Botrytis virus F, a flexuous rod-shaped mycovirus resembling plant 'potex-like' viruses.";	
RL J. Gen. Virol. 82:67-78(2001).	
RN [2]	
RP SEQUENCE FROM N.A.	
RA Howitt R.L.J., Beaver R.E., Pearson M.N., Forster R.L.S.;	
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AF238884; AAC23416.1;	
DR InterPro: IPR002965; P_Rich_extensn.	
DR InterPro: IPR006066; Viral_helicesn.	
DR Pfam: PF01443; Viral_helicesn.1.	
DR PRINTS: PR01217; PRICHEXTENS.	
SQ SEQUENCE 1896 AA; 212376 MW; EE15A3BD1387B271 CRC64;	
Query Match	42.9%; Score 78.5; DB 12; Length 1896;
Best Local Similarity	52.0%; Pred. No. 0.4;
Matches 13; Conservative 2; Mismatches 7; Indels 3; Gaps 1;	
QY 5 WPKWPKWPLIGG---YDPAPPPPP 26	
DB 817 WPKWPKWPLIGG---YDPAPPPPP 841	
RESULT 2	
Q08198	PRELIMINARY; PRT; 161 AA.
ID Q08198	

AC Q08198;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE CYSTEINE-RICH EXTENSIN-LIKE PROTEIN.  
 OS Nicotiana tabacum (Common tobacco).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93342083; PubMed=83411705;  
 RA Wu H.M., Zou J., May B., Gu Q., Cheung A.Y.;  
 RT "A tobacco gene family for flower cell wall proteins with a proline-  
 rich domain and a cysteine-rich domain.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6829-6833(1993).  
 DR EMBL; I13443; AAA34063.1;  
 DR Mendel; 16905; Nicita;2747;16905.  
 SQ SEQUENCE 161 AA; 17791 MW; E3FD267EA21A2C66 CRC64;

Query Match  
 Best Local Similarity 42.1%; Score 77; DB 10; Length 161;  
 Matches 14; Conservative 2; Mismatches 2; Indels 14; Gaps 3;

QY 3 WPM-----WPM--KWPLGGGVDAPPAPPP 26  
 DB 41 WPMKIPCYLMPFPPWPC-----PSPPPPP 66

RESULT 3  
 ID 066118 PRELIMINARY; PRT; 239 AA.  
 AC 066118;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)  
 DE HYPOTHETICAL 26.2 KDA PROTEIN.  
 OS Zymomonas mobilis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;  
 OC Zymomonas.  
 OX NCBI\_TaxID=542;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=ATCC 29191.  
 RX MEDLINE=98175679; PubMed=9515924;  
 RA Neveling U., Klaseen R., Bringer-Meyer S., Sahm H.;  
 RT "Purification of the pyruvate dehydrogenase multienzyme complex of  
 RT Zymomonas mobilis and identification and sequence analysis of the  
 RT corresponding genes.";  
 DR EMBL; X93605; CAA63807.1;  
 DR J. Bacteriol. 180:1540-1548(1998).  
 KW Hypothetical protein.  
 SQ SEQUENCE 239 AA; 26249 MW; 08698FCA68085FDE CRC64;

Query Match  
 Best Local Similarity 39.3%; Score 72; DB 2; Length 239;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 14 GGGYDPAPPPPP 26  
 DB 176 GGAAYQAPAPPPPP 188

RESULT 4  
 ID 023248 PRELIMINARY; PRT; 172 AA.  
 AC 023248;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2000 (TREMBLrel. 14, last annotation update)

DE ZC168.5 PROTEIN.  
 GN ZC168.5.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Peloderae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Berks M.;  
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightings J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).  
 DR EMBL; Z70312; CAA94385.1;  
 SQ SEQUENCE 172 AA; 18121 MW; 67276F9096DBC8F CRC64;

Query Match  
 Best Local Similarity 38.3%; Score 70; DB 5; Length 172;  
 Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 11 PLIGGYDPAPPPPP 26  
 DB 45 PPGCGGYEAPPPPP 60

RESULT 5  
 ID 061078 PRELIMINARY; PRT; 520 AA.  
 AC 061078;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)  
 DE WISCOTT-ALDRICH SYNDROME PROTEIN HOMOLOG.  
 GN N-3APL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=NIH-SWISS;  
 RA Li W., Margolis B., Schlessinger J.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U42471; AA85515.1;  
 DR InterPro; IPR000095; PAK\_box\_P21\_Rho\_binding.  
 DR InterPro; IPR000697; RanBP1\_WASP.  
 DR InterPro; IPR003124; WH2.  
 DR Pfam; PF00568; WH1; 1.  
 DR Pfam; PF00786; PBD; 1.  
 DR Pfam; PF02205; WH2; 1.  
 DR PROSITE; PS50108; GBD; 1.  
 DR SMART; SM00285; PBD; 1.  
 DR SMART; SM00461; WH1; 1.  
 DR SMART; SM00246; WH2; 1.  
 SQ SEQUENCE 520 AA; 54260 MW; 9880213DE59524B2 CRC64;

Query Match  
 Best Local Similarity 37.7%; Score 69; DB 11; Length 520;

Best Local Similarity 75.0%; Pred. No. 1.5;  
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 11 PLIGGYDAPPPPP 26  
DB 401 PLGAGGPPPPPP 416

RESULT 6  
049946 PRELIMINARY; PRT; 221 AA.

AC 049946; TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_Taxid=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. RECORD: TISSUE-SMELLING STOLON;  
RA MacLeod M.R., Taylor M.A., Davies H.V.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ003220; CAA06000.1;  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR PRINTS; PR01217; PRICHEXTNSN.  
FT NON\_TER 1  
SQ SEQUENCE 221 AA; 24382 MW; F30F469B60727335 CRC64;

Query Match 37.4%; Score 68.5; DB 10; Length 221;  
Best Local Similarity 36.8%; Pred. No. 0.74;  
Matches 14; Conservative 3; Mismatches 4; Indels 17; Gaps 3;

OY 3 WPM-----WPMKWPILGGYD-----PAPPPPP 26  
DB 24 WPMELPCYLPMPFPF---RPPCPPPKPPSPPPPP 58

RESULT 7  
040548 PRELIMINARY; PRT; 109 AA.

AC 040548; TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE PISTIL-SPECIFIC EXTENSIN-LIKE PROTEIN PRECURSOR (FRAGMENT).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_Taxid=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. PETITE HAVANA; TISSUE-PISTIL;  
RA MEDLINE-93005740; PubMed-1392607;  
RA Goldman S., Pezzotti M., Seurinck J., Mariani C.;  
RT "Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins."  
RL Plant Cell 4:1041-1051(1992).  
CC -1- TISSUE SPECIFICITY: PISTIL (STIGMA AND STYLE TISSUE).  
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS IN FLORAL BUDS AFTER PISTIL  
CC DIFFERENTIATION AND LEVELS GRADUALLY INCREASE DURING FLOWER  
CC DEVELOPMENT TOWARD ANTHESIS. LEVELS GRADUALLY DECREASE AFTER  
CC POLLINATION AND ARE ABSENT BY THE SIXTH DAY AFTER POLLINATION.  
EMBL; Z14014; CAA78392.1;  
DR Mendel; 16906; Nicta; 2747; 16906.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR PRINTS; PR01217; PRICHEXTNSN.  
KW structural protein; Repeat; Glycoprotein; Signal.

FT SIGNAL 1 22  
FT CHAIN 23 109  
FT DOMAIN 81 109  
FT REPEAT 81 85  
FT REPEAT 93 97  
FT REPEAT 105 109  
FT NON\_TER 109 109  
SQ SEQUENCE 109 AA; 12373 MW; 6A6BCEFA192ABC8 CRC64;

Query Match 37.2%; Score 68; DB 10; Length 109;  
Best Local Similarity 36.7%; Pred. No. 0.42;  
Matches 12; Conservative 1; Mismatches 2; Indels 16; Gaps 2;

OY 3 WPM-----WPMKWPILGGYDAPPPPP 26  
DB 43 WPMELPCYLPMPFPF-----PPPPPP 64

RESULT 8

ID 09P3K2 PRELIMINARY; PRT; 142 AA.  
AC 09P3K2;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE RELATED TO EXTENSIN PRECURSOR.  
GN B15120.90.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_Taxid=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Schulte U., Algen V., Hobeisel J., Brandt F., Fartmann B., Holland R.,  
RA Makatura G., Mewes H.W., Mannhaupt G.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
RP SEQUENCE FROM N.A.  
RA German Neurospora genome project;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL389900; CAB97463.1;  
SQ SEQUENCE 142 AA; 15793 MW; C1423229C79068FC CRC64;

Query Match 37.2%; Score 68; DB 3; Length 142;  
Best Local Similarity 42.9%; Pred. No. 0.55;  
Matches 15; Conservative 3; Mismatches 7; Indels 10; Gaps 2;

OY 1 RRP-----WPMKWPILG---GGYDAPPPPP 25  
DB 87 RRPPLTRGSGWPPKPPPLTSFLLSFGAPPPPP 121

RESULT 9

ID 09FYZ6 PRELIMINARY; PRT; 150 AA.  
AC 09FYZ6;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE EIN3 (FRAGMENT).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
OX NCBI\_Taxid=4097;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV. KY57;  
RA Terajima Y., Satoh S.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF247568; AAG00419.1;

FT NON\_TER 1 1  
RT NON\_TER 150 150  
SQ SEQUENCE 150 AA; 16854 MW; 126710F8925140A7 CRC64;

Query Match  
Best local Similarity 37.2%; Score 68; DB 10; Length 150;  
Matches 12; Conservative 1; Mismatches 6; Indels 6; Gaps 2;

OY 4 PWP-----WKWPLIGGYDPAPP 23  
DB 118 PWPPTGQEDW-WPOLGLSKDQGP 141

RESULT 10  
ID 008195 PRELIMINARY; PRT; 196 AA.

AC 008195;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, last sequence update)  
DT 01-JUN-2000 (TREMblrel. 14, last annotation update)  
DE CYSTEINE-RICH EXTENSIN-LIKE PROTEIN 2.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
ON NCBI\_TaxID=4097;  
[1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=93342083; Pubmed=8341705;  
RA Wu H.M., Zou J., May B., Gu Q., Cheung A.Y.;  
RT "A tobacco gene family for flower cell wall proteins with a proline-  
rich domain and a cysteine-rich domain."  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6829-6833(1993).  
DR EMBL: L13440; AAA34060.1;  
DR Mendel: 16902; Nicita:2747;16902.  
SQ SEQUENCE 196 AA; 21913 MW; 4C44E23C5B706E30 CRC64;

Query Match  
Best local Similarity 37.2%; Score 68; DB 10; Length 196;  
Matches 12; Conservative 1; Mismatches 2; Indels 16; Gaps 2;

OY 3 WPM-----WPMKMPILGGYDPAPP 26  
DB 40 WPMELPCYLTWPFPP-----PPPP 61

RESULT 11  
ID 008194 PRELIMINARY; PRT; 209 AA.

AC 008194;  
DT 01-JAN-1998 (TREMblrel. 05, Created)  
DT 01-JAN-1998 (TREMblrel. 05, last sequence update)  
DT 01-JUN-2000 (TREMblrel. 14, last annotation update)  
DE PISTIL-SPECIFIC EXTENSIN-LIKE PROTEIN PRECURSOR  
DE (CYSTEINE-RICH EXTENSIN-LIKE PROTEIN 1).  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
ON NCBI\_TaxID=4097;  
[1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=93342083; Pubmed=8341705;  
RA Wu H.M., Zou J., May B., Gu Q., Cheung A.Y.;  
RT "A tobacco gene family for flower cell wall proteins with a proline-  
rich domain and a cysteine-rich domain."  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6829-6833(1993).  
DR EMBL: L13440; AAA34060.1;  
DR Mendel: 16902; Nicita:2747;16902.  
SQ SEQUENCE 209 AA; 23351 MW; 12198BE2B8E08ED5 CRC64;

RP SEQUENCE OF 39-209 FROM N.A.  
RX STRAIN=CV. PETITE HAVANA; TISSUE=PISTIL;  
RX MEDLINE=93005740; Pubmed=1392607;

RA Goldman S., Dezzotti M., Saurinck J., Mariani C.;  
RT "Developmental expression of tobacco pistil-specific genes encoding  
novel extensin-like proteins."  
RL Plant Cell 4:1041-1051(1992).

CC -1- TISSUE SPECIFICITY: PISTIL (STIGMA AND STYLE TISSUE).  
CC -1- DEVELOPMENTAL STAGE: EXPRESSION BEGINS IN FLORAL BUDS AFTER PISTIL  
DIFFERENTIATION AND LEVELS GRADUALLY INCREASE DURING FLOWER  
DEVELOPMENT TOWARD ANTHESIS. LEVELS GRADUALLY DECREASE AFTER  
POLLINATION AND ARE ABSENT BY THE SIXTH DAY AFTER POLLINATION.  
CC POLININATION AND ARE ABSENT BY THE SIXTH DAY AFTER POLLINATION.

DR EMBL: L13439; AAA34059.1;  
DR EMBL: 214020; CAA78398.1;  
DR Mendel: 16901; Nicita:2747;16901.

KW Structural protein; Repeat; Glycoprotein; Signal.  
FT SIGNAL 1 19  
FT CHAIN 20 209 PISTIL-SPECIFIC EXTENSIN-LIKE PROTEIN.  
FT DOMAIN 81 126 5 X 5 AA REPEATS OF S-P(4).  
FT REPEAT 81 85 1.  
FT REPEAT 93 97 2.  
FT REPEAT 105 109 3.  
FT REPEAT 117 121 4.  
FT REPEAT 122 126 5.  
FT CARBOHYD 146 146 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 209 AA; 23351 MW; 12198BE2B8E08ED5 CRC64;

Query Match  
Best local Similarity 37.2%; Score 68; DB 10; Length 209;  
Matches 12; Conservative 1; Mismatches 2; Indels 16; Gaps 2;

OY 3 WPM-----WPMKMPILGGYDPAPP 26  
DB 45 WPMELPCYLTWPFPP-----PPPP 66

RESULT 12  
ID 092WK1 PRELIMINARY; PRT; 615 AA.

AC 092WK1;  
DT 01-MAY-1999 (TREMblrel. 10, Created)  
DT 01-MAY-1999 (TREMblrel. 10, last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
DE TRANSCRIPTION FACTOR TEIL.  
OS Nicotiana tabacum (Common tobacco).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
ON NCBI\_TaxID=4097;  
[1]

RP SEQUENCE FROM N.A.  
RA Kosugi S., Ohashi Y.;  
RT "A ETHYLENE-INSENSITIVE3 homolog from tobacco is a transcription  
factor with sequence-specific DNA binding and transactivating  
functions."  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB015855; CAB19436.1;  
DR Mendel: 38225; Nicita:3163;38225.  
DR InterPro: IPR001064; CysTallin.  
DR PROSITE: PS00225; CRYSTALLIN-BETAGAMMA; UNKNOWN 1.  
SQ SEQUENCE 615 AA; 69969 MW; 20F8F9774A259D2C CRC64;

Query Match  
Best local Similarity 37.2%; Score 68; DB 10; Length 615;  
Matches 12; Conservative 1; Mismatches 6; Indels 6; Gaps 2;

OY 4 PWP-----WKWPLIGGYDPAPP 23  
DB 215 PWPPTGQEDW-WPOLGLSKDQGP 238

RESULT 13  
ID 017795 PRELIMINARY; PRT; 969 AA.

AC Q17795;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE SIMILARITY TO WISCOTT-ALDRICH SYNDROME PROTEINS.  
 GN C07G1.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
 RA Smailton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans."  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX Hawkins J.;  
 RN [3]  
 RP Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RA Waterston R.;  
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; U58751; AAB0657.1; -;  
 DR InterPro: IPR000095; PAK\_box\_P21\_Rho-binding.  
 DR InterPro: IPR000697; RanBP1\_WASP.  
 DR InterPro: IPR001960; WH1.  
 DR InterPro: IPR003124; WH2.  
 DR Pfam: PF00568; WH1; 1.  
 DR Pfam: PF00786; PBD; 1.  
 DR Pfam: PF02205; WH2; 2.  
 DR PROSITE: PS50108; GBD; 1.  
 DR SMART: SM00285; PBD; 1.  
 DR SMART: SM00461; WH1; 1.  
 DR SMART: SM00246; WH2; 2.  
 SO SEQUENCE 969 AA; 105790 MW; BE2D115D0C620BE8 CRC64;  
 Query Match 37.2%; Score 68; DB 5; Length 969;  
 Best Local Similarity 55.0%; Pred. No. 3.6;  
 Matches 11; Conservative 1; Mismatches 8; Indels 0; Gaps 0;  
 QY 7 PWKWLIGGGYDPAPPPPP 26  
 Db 799 PMGLPAVGAGAPPPPPPP 818  
 RESULT 14  
 ID 094901 PRELIMINARY; PRT: 352 AA.  
 AC 094901;  
 DT 01-FEB-1997 (Tremblrel. 02, Created)  
 DT 01-FEB-1997 (Tremblrel. 02, last sequence update)  
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
 DE LARK PROTEIN.  
 GN LARK OR CG8597.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyrtoidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Ceinikher S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Chertey J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaes R., Tector C., Turner R., Venter E., Wang A., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RL "The genome sequence of Drosophila melanogaster."  
 RT Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Newby L.M., Jackson F.R.;  
 RL J. Neurobiol. 13:0-0(0).  
 CC -1- SIMILARITY: BELONGS TO ZN-FINGER CCHC TYPE FAMILY.  
 DR EMBL; AE003559; AAF50578.1; -;  
 DR EMBL; U59476; AAB07067.1; -;  
 DR FLYbase: FBgn0011640; lark.  
 DR InterPro: IPR001878; Znf\_CCHC.  
 DR InterPro: IPR000504; RRM.  
 DR InterPro: IPR001076; RRM; 2.  
 DR Pfam: PF00098; zf-CCHC; 1.  
 DR Pfam: PF00098; zf-CCHC; 1.  
 DR PRINTS: PR00939; C2HCFINGER.  
 DR SMART: SM00360; RRM; 2.  
 DR SMART: SM00343; ZNF\_C2HC; 1.  
 DR PROSITE: PS50102; RRM; 2.  
 KW Zinc-finger.  
 SO SEQUENCE 352 AA; 39912 MW; C8D698D4DD122FA9 CRC64;  
 Query Match 36.9%; Score 67.5; DB 5; Length 352;  
 Best Local Similarity 56.5%; Pred. No. 1.5;  
 Matches 13; Conservative 0; Mismatches 3; Indels 7; Gaps 1;  
 QY 11 PLTGGY-----DPAPPPPP 26  
 Db 199 PLTAGGYRDMRGDPYPPPPPP 221

RESULT 15

ID 008197 PRELIMINARY; PRT; 157 AA.  
 AC 008197;  
 DT 01-NOV-1996 (TRENBLREL. 01, Created)  
 DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
 DT 01-JUN-2000 (TRENBLREL. 14, Last annotation update)  
 DE CYSTEINE-RICH EXTENSIN-LIKE PROTEIN 4.  
 OS Nicotiana tabacum (Common tobacco);  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.  
 OX NCBI\_TaxID=4097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93342083; PubMed=8341705;  
 RA Wu H.M., Zou J., Gu Q., Cheung A.Y.;  
 RT "A tobacco gene family for flower cell wall proteins with a proline-  
 rich domain and a cysteine-rich domain."  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6829-6833(1993).  
 DR EMBL: U3442; AAA34062.1;  
 DR Mende1; 16904; Nicta; 2747; 16904.  
 DR Mende1; 157 AA; 17459 MW; 6EBA02249CF6A35 CRC64;  
 SQ SEQUENCE

Query Match 36.3%; Score 66.5; DB 10; Length 157;  
 Best local Similarity 35.0%; Pred. No. 0.91;  
 Matches 14; Conservative 2; Mismatches 5; Indels 19; Gaps 3;

OY 3 WPM-----WPMKPLIGGYD-----PAPPPPP 26  
 ||| ||:| | | |||||  
 Db 36 WPMKIPCYLRWPFPP---RPYPCSPPRPRPCPPPPPP 72

Search completed: January 30, 2002, 11:52:00  
 Job time: 199 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2002, 11:52:22 ; Search time 18.17 Seconds  
(without alignments)  
28.250 Million cell updates/sec

Title: US-09-432-546-5  
Perfect score: 103  
Sequence: 1 SRRWPMWPKWPLI 14

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	60.2	144	1	INDC_BOVIN
2	53.5	51.9	491	1	ADRO_HUMAN
3	52	50.5	1112	1	CN3B_HUMAN
4	51.5	50.0	278	1	RCEL_RHOG
5	51.5	50.0	492	1	ADRO_BOVIN
6	51	49.5	1173	1	VGL2_CVR22
7	50.5	49.0	494	1	ADRO_RAT
8	49.5	48.1	253	1	Y945_MYCTU
9	48.5	47.1	212	1	PF11_PIG
10	48.5	47.1	228	1	PF12_PIG
11	48	46.6	990	1	ENV_OMVVS
12	47	45.6	942	1	ENV_CAEVC
13	47	45.6	966	1	ENV_CAEVC
14	47	45.6	982	1	ENV_VILV
15	47	45.6	983	1	ENV_VILV
16	47	45.6	989	1	ENV_VILV1
17	47	45.6	991	1	ENV_VILV2
18	46	44.7	533	1	UBIB_PSEAE
19	46	44.7	572	1	GAB_IPHA
20	45.5	44.2	660	1	FHUB_ECOLI
21	45	43.7	115	1	NUFM_BOVIN
22	45	43.7	115	1	NUFM_RAT
23	45	43.7	214	1	AVEN_AVESA
24	45	43.7	534	1	APG_ARATH
25	45	43.7	722	1	ACM1_DROME
26	44	42.7	115	1	NUFM_HUMAN
27	44	42.7	407	1	FUCO_CAEEL
28	44	42.7	485	1	AER5_AERHY
29	44	42.7	488	1	AERA_AERSO
30	44	42.7	489	1	AERA_AERSA
31	44	42.7	492	1	AER3_AERHY
32	44	42.7	492	1	AER4_AERHY
33	44	42.7	492	1	AERA_AERTR

34	44	42.7	493	1	AERA_AERHY
35	44	42.7	543	1	CP1B_RAT
36	44	42.7	655	1	HGFA_HUMAN
37	44	42.7	984	1	SV13_MOUSE
38	43.5	42.2	515	1	ENV_BLV
39	43.5	42.2	515	1	ENV_BLVAF
40	43.5	42.2	515	1	ENV_BLVAV
41	43.5	42.2	515	1	ENV_BLVAV
42	43.5	42.2	515	1	ENV_BLVAV
43	43.5	42.2	515	1	ENV_BLVAV
44	43.5	42.2	515	1	ENV_BLVAV
45	43	41.7	255	1	RCEL_RHOPH

## ALIGNMENTS

RESULT 1	
INDC_BOVIN	
ID	INDC_BOVIN
AC	P33046;
DT	01-OCT-1993 (Rel. 27, Created)
DT	01-OCT-1993 (Rel. 27, Last sequence update)
DT	01-NOV-1997 (Rel. 35, Last annotation update)
DE	INDOLICIDIN PRECURSOR.
OS	Bos taurus (Bovine)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC	Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Bone marrow;
RX	MEDLINE=92392368; PubMed=1520337;
RA	del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;
RT	"cDNA cloning of the neutrophil bactericidal peptide indolicidin.";
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).
RN	[2]
RP	SEQUENCE OF 131-143.
RC	TISSUE=Neutrophils;
RX	MEDLINE=92165771; PubMed=1537821;
RA	Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.,
RT	Cullor J.S.;
RT	"Indolicidin, a novel bactericidal tridecapeptide amide from
RT	neutrophils.";
RL	J. Biol. Chem. 267:4292-4295(1992).
CC	-!- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST
CC	STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.
CC	-!- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.
CC	-!- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.
CC	-!- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
CC	EMBL; X67340; CAA47755.1;
DR	PIR; JCI222; JCI222.
DR	PIR; A42387; A42387.
DR	InterPro; IPR001894; Cathelicidin.
DR	Pfam; PF00666; Cathelicidins; 1.
DR	PROSITE; PS001838; Cathelicidin; 1.
DR	PROSITE; PS00946; CATHELICIDINS_1; 1.
DR	PROSITE; PS00947; CATHELICIDINS_2; 1.
KW	Antibiotic; Amidation; Signal.
FT	SIGNAL 1 29
FT	PROPEP 30 130
FT	PEPTIDE 131 143
FT	MOD_RES 30 30
FT	INDOLICIDIN.
FT	PYRROLIDONE CARBOXYLIC ACID (BY

Thu Jan 31 11:07:41 2002

FT DISULFID 85 96  
 FT DISULFID 107 124  
 FT MOD\_RES 143 143  
 SQ SEQUENCE 144 AA; 16479 MW; E3B1CB55C09911 CRC64;  
 SIMILARITY).  
 BY SIMILARITY.  
 BY SIMILARITY.  
 AMIDATION (G-144 PROVIDE AMIDE GROUP).  
 Query Match 60.28; Score 62; DB 1; Length 144;  
 Best Local Similarity 75.04; Pred. No. 0.088; 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 2; Mismatches 0;  
 QY 3 RWPWPWK 10  
 DB 135 KWPWPWR 142  
 RESULT 2  
 ID ADRO\_HUMAN STANDARD; PRT; 491 AA.  
 AC P22570; Q13716;  
 DT 01-AUG-1991 (Rel. 19, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE NADPH:ADRENODOXIN OXIDOREDUCTASE, MITOCHONDRIAL PRECURSOR  
 (EC 1.18.1.2) (ADRENODOXIN REDUCTASE) (AR) (FERREDOXIN-NADP(+)  
 REDUCTASE).  
 GN FOXR OR ADXR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=89017146; PubMed=2845336;  
 RA Solish S.B., Picado-Leonard J., Morel Y., Kuhn R.W., Mohandas T.K.,  
 Hanukoglu I., Miller W.L.;  
 RT "Human adrenodoxin reductase: two mRNAs encoded by a single gene on  
 chromosome 17cen--q25 are expressed in steroidogenic tissues.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:7104-7108(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91046028; PubMed=2236061;  
 RA Lin D., Shi Y., Miller W.L.;  
 RT "Cloning and sequence of the human adrenodoxin reductase gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8516-8520(1990).  
 CC -1- FUNCTION: SERVES AS THE FIRST ELECTRON TRANSFER PROTEIN IN ALL THE  
 MITOCHONDRIAL P450 SYSTEMS. INCLUDING CHOLESTEROL SIDE CHAIN  
 CLEAVAGE IN ALL STEROIDGENIC TISSUES. STEROID 11-BETA  
 HYDROXYLATION IN THE ADRENAL CORTEX, 25-OH-VITAMIN D3-24  
 HYDROXYLATION IN THE KIDNEY, AND STEROL C-27 HYDROXYLATION IN THE  
 LIVER.  
 CC -1- CATALYTIC ACTIVITY: REDUCED ADRENODOXIN + NADP(+) = OXIDIZED  
 ADRENODOXIN + NADPH.  
 CC -1- COFACTOR: FAD.  
 CC -1- PATHWAY: CHOLESTEROL SIDE-CHAIN-CLEAVAGE SYSTEM.  
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A  
 LONG FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE LONG FORM  
 REPRESENTS 10-20% OF ALL ADRENODOXIN REDUCTASE MRNA. AND SEEMS TO  
 BE INACTIVE.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL; J03826; AAB59498.1;  
 DR EMBL; J03826; AAB59497.1;  
 DR EMBL; M58509; AAA51668.1;  
 DR EMBL; M58508; AAA51668.1; JOINED.

DR EMBL; M58509; AAA51669.1;  
 DR EMBL; M58508; AAA51669.1; JOINED.  
 DR PIR; A36482; A36482.  
 DR PIR; A40487; A40487.  
 DR MIW; 103270;  
 DR InterPro: IPR000759; Adrdx\_redctse.  
 DR PRINTS; PR00419; ADXRDTASE.  
 KW Electron transport; Oxidoreductase; Flavoprotein; NADP; FAD;  
 Mitochondrion; Transit peptide; Alternative splicing; Polymorphism.  
 FT TRANSIT 33 491  
 FT CHAIN 203 203  
 FT VARSPLIC 123 123  
 FT VARIANT 123 123  
 FT Q -> R.  
 FT /FTID=VAR\_004624.  
 FT SEQUENCE 491 AA; 53808 MW; 85865BAA2276D2B2 CRC64;  
 Query Match 51.98; Score 53.5; DB 1; Length 491;  
 Best Local Similarity 61.5%; Pred. No. 3.1;  
 Matches 8; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 QY 1 SRRWPWPWK-WP 12  
 DB 3 SRCWRWGSAMP 15  
 RESULT 3  
 ID CN3B\_HUMAN STANDARD; PRT; 1112 AA.  
 AC Q13370; O00639; Q14408;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE CGMP-INHIBITED 3',5'-CYCLIC PHOSPHODIESTERASE B (EC 3.1.4.17) (CYCLIC  
 GMP INHIBITED PHOSPHODIESTERASE B) (CGI-PDE B) (CGIPDE1) (CGIP1).  
 GN PDE3B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Fat;  
 MEDLINE=97038690; PubMed=8884271;  
 RA Miki T., Taira M., Hockman S., Shimada F., Lieman J., Napolitano M.,  
 Ward D., Taira M., Makino H., Manganiello V.C.;  
 RT "Characterization of the cDNA and gene encoding human PDE3B, the  
 CGI1 isoform of the human cyclic GMP-inhibited cyclic nucleotide  
 phosphodiesterase family.";  
 RL Genomics 36:476-485(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97079687; PubMed=8921398;  
 RA Loeber R.W., Winterpacht A., Seipel B., Zabel B.U.;  
 RT "Molecular cloning and chromosomal assignment of the human homologue  
 of the rat cGMP-inhibited phosphodiesterase 1 (PDE3A) -- a gene  
 involved in fat metabolism located at 11p15.1.";  
 RL Genomics 37:211-218(1996).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN FAT METABOLISM.  
 CC -1- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O -  
 GUANOSINE 5'-PHOSPHATE.  
 CC -1- ENZYME REGULATION: INHIBITED BY CGMP.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).  
 CC -1- TISSUE SPECIFICITY: ABUNDANT IN ADIPOSE TISSUES.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE  
 FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 DR EMBL; J03826; AAB59498.1;  
 DR EMBL; J03826; AAB59497.1;  
 DR EMBL; M58509; AAA51668.1;  
 DR EMBL; M58508; AAA51668.1; JOINED.



or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```

CC  EMBL; U38178; AAC50724.1; -
CC  EMBL; D50640; BAA09306.1; -
CC  EMBL; D50625; BAA09306.1; JOINED.
CC  EMBL; D50626; BAA09306.1; JOINED.
CC  EMBL; D50627; BAA09306.1; JOINED.
CC  EMBL; D50628; BAA09306.1; JOINED.
CC  EMBL; D50629; BAA09306.1; JOINED.
CC  EMBL; D50630; BAA09306.1; JOINED.
CC  EMBL; D50631; BAA09306.1; JOINED.
CC  EMBL; D50632; BAA09306.1; JOINED.
CC  EMBL; D50633; BAA09306.1; JOINED.
CC  EMBL; D50634; BAA09306.1; JOINED.
CC  EMBL; D50635; BAA09306.1; JOINED.
CC  EMBL; D50637; BAA09306.1; JOINED.
CC  EMBL; D50638; BAA09306.1; JOINED.
CC  EMBL; D50639; BAA09306.1; JOINED.
CC  EMBL; X95520; CAA64774.1; -
CC  MIN; 602047; -
CC  InterPro; IPR003607; HDC.
CC  InterPro; IPR002073; PDEase.
CC  Pfam; PF00233; PDEase; 1.
CC  SMART; SM00471; HDC; 1.
CC  PROSITE; PS00126; PDEASE_I; 1.
CC  Hydrolase; CGMP; Membrane.
CC  DOMAIN; 1077 1080
CC  CONFLICT; 84 84
CC  CONFLICT; 87 87
CC  SEQUENCE; 1112 AA; 124376 MW; E5141C3DA12E99B0 CRC64;

Query Match 50.5%; Score 52; DB 1; Length 1112;
Best Local Similarity 62.5%; Pred. No. 9.9;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 WPWPWKW 11
    | | | |
Db 169 WQWNSWPW 176

RESULT 4
RCBL_RHOGE
ID RCBL_RHOGE STANDARD; PRT; 278 AA.
AC P51760;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE REACTION CENTER PROTEIN L CHAIN (PHOTOSYNTHETIC REACTION CENTER L
DE SUBUNIT).
GN PUFL.
OS Rhodocyclus gelatinosus (Rhodospseudomonas gelatinosa).
OC Bacteria; Proteobacteria; beta subdivision; Comamonadaceae;
OC Rubrivivax.
OX NCBI_TaxID=28068;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IL144;
RX MEDLINE=94132007; PubMed=8300574;
RA Nagashima K.V.P., Matsura K., Ohya S., Shimada K.;
RT "Primary structure and transcription of genes encoding B870 and
RT photosynthetic reaction center apoproteins from Rubrivivax
RT gelatinosus."
RL J. Biol. Chem. 269:2477-2484(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SI;
RX MEDLINE=97033541; PubMed=8879238;
RA Ouchane S., Picard M., Reiss-Husson F., Vernotte C., Astier C.;
RT "Development of gene transfer methods for Rubrivivax gelatinosus S1:
RT construction, characterization and complementation of a puf operon
RT deletion strain."
RL Mol. Gen. Genet. 252:379-385(1996).

```

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

```

CC  EMBL; D16822; BAA04100.1; -
CC  EMBL; U30310; AAA73927.1; -
CC  EMBL; U51298; ABA11576.1; -
CC  HSP; P02954; 2RCR.
CC  InterPro; IPR000484; Photo_RC.
CC  Pfam; PF00124; photoRC; 1.
CC  PRINTS; PR00256; REACTNCTR.
CC  PRODOM; PD000551; Photo_RC; 1.
CC  PROSITE; PS00244; REACTION_CENTER; 1.
CC  Transmembrane; Electron transport; Photosynthesis; Reaction center;
CC  Iron; Magnesium.
CC  INIT_MET 0 0 BY SIMILARITY.
CC  TRANSMEM 32 55 POTENTIAL.
CC  TRANSMEM 84 112 POTENTIAL.
CC  TRANSMEM 115 140 POTENTIAL.
CC  TRANSMEM 170 199 POTENTIAL.
CC  TRANSMEM 225 251 POTENTIAL.
CC  METAL 173 173 MAGNESIUM (BACTERIOCHLOROPHYLL B) (BY
CC  SIMILARITY).
CC  METAL 190 190 IRON (NON HAEM) (BY SIMILARITY).
CC  METAL 230 230 IRON (NON HAEM) (BY SIMILARITY).
CC  BINDING 216 216 QUINONE B (BY SIMILARITY).
CC  CONFLICT 278 278 K -> N (IN REF. 2).
CC  SEQUENCE 278 AA; 31210 MW; 0ACCFC1241890DC5 CRC64;

Query Match 50.0%; Score 51.5; DB 1; Length 278;
Best Local Similarity 40.9%; Pred. No. 3.3;
Matches 9; Conservative 2; Mismatches 2; Indels 9; Gaps 2;

QY 1 SRWP-WPWP-----KWPL 13
    | | | | | | | | | |
Db 256 TRGWPEWGWNLNLPWSQWPL 277

RESULT 5
ADRO_BOVIN
ID ADRO_BOVIN STANDARD; PRT; 492 AA.
AC P08165;
DT 01-AUG-1988 (Rel. 08, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NADPH:ADRENODOXIN OXIDOREDUCTASE, MITOCHONDRIAL PRECURSOR
DE (EC 1.18.1.2) (ADRENODOXIN REDUCTASE) (AR) (FERREDOXIN-NADP(+))
DE REDUCTASE).
GN FDXR OR ADXR
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
RX MEDLINE=94177140; PubMed=8130767;
RA Takata Y., Sagara Y., Kono A., Sekimizu K., Horiuchi T.;

```

Thu Jan 31 11:07:41 2002

us-09-432-546-5.rsp

RT "Gene structure of bovine adrenodoxin reductase.";  
 RL Biol. Pharm. Bull. 16:1200-1206(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=88198050; PubMed=3448086;  
 RA Sagara Y., Takata Y., Miyata T., Hara T., Horiuchi T.;  
 RT Cloning and sequence analysis of adrenodoxin reductase cDNA from  
 RT bovine adrenal cortex.";  
 RL J. Biochem. 102:1333-1336(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87270696; PubMed=3038094;  
 RA Nonaka Y., Okamoto H., Yabusaki Y., Kuramitsu S., Kagamiyama H.,  
 RA Yamano T., Okamoto M.;  
 RT Molecular cloning and sequence analysis of full-length cDNA for mRNA  
 RT of adrenodoxin oxidoreductase from bovine adrenal cortex.";  
 RL Biochem. Biophys. Res. Commun. 145:1239-1247(1987).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Adrenal cortex;  
 MEDLINE=89170752; PubMed=2924777;  
 RA Hanukoglu I., Gutfinger T.;  
 RT "cDNA sequence of adrenodoxin reductase. Identification of NADP-  
 RT binding sites in oxidoreductases.";  
 RL Eur. J. Biochem. 180:479-484(1989).  
 RN [5]  
 RP SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.  
 RX TISSUE=Adrenal cortex;  
 MEDLINE=88082777; PubMed=3691502;  
 RA Hanukoglu I., Gutfinger T., Haniu M., Shively J.E.;  
 RT "Isolation of a cDNA for adrenodoxin reductase (ferredoxin-NADP+  
 RT reductase). Implications for mitochondrial cytochrome P-450 systems.";  
 RL Eur. J. Biochem. 169:449-455(1987).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 33-492.  
 RX TISSUE=Adrenal gland;  
 MEDLINE=99299392; PubMed=10369776;  
 RA Ziegler G.A., Vohrhehn C., Hanukoglu I., Schulz G.E.;  
 RT "The structure of adrenodoxin reductase (ferredoxin-NADP+  
 RT electron transfer for steroid biosynthesis.";  
 RL J. Mol. Biol. 289:981-990(1999).  
 CC -!- FUNCTION: SERVES AS THE FIRST ELECTRON TRANSFER PROTEIN IN ALL THE  
 CC MITOCHONDRIAL P450 SYSTEMS. INCLUDING CHOLESTEROL SIDE CHAIN  
 CC CLEAVAGE IN ALL STEROID-GENIC TISSUES, STEROID 11-BETA  
 CC HYDROXYLATION IN THE ADRENAL CORTEX. 25-OH-VITAMIN D3-24  
 CC HYDROXYLATION IN THE KIDNEY, AND STEROL C-27 HYDROXYLATION IN THE  
 CC LIVER.  
 CC -!- CATALYTIC ACTIVITY: REDUCED ADRENODOXIN + NADP(+) = OXIDIZED  
 CC ADRENODOXIN + NADPH.  
 CC -!- COFACTOR: FAD.  
 CC -!- PATHWAY: CHOLESTEROL SIDE-CHAIN-CLEAVAGE SYSTEM.  
 CC -!- SUBUNIT: MONOMER.  
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.  
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A  
 CC LONG FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE LONG FORM  
 CC REPRESENTS 10-20% OF ALL ADRENODOXIN REDUCTASE MRNA. AND SEEMS TO  
 CC BE INACTIVE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M17029; AAA30362.1;  
 CC EMBL: D00211; BAA00150.1;  
 CC EMBL: X13736; CAA32002.1;  
 CC PIR: A29604; A29604.  
 CC PIR: J50390; J50390.  
 CC PIR: S03558; S03558.  
 CC PIR: J70751; J70751.  
 DR PDB: 1CJC; 12-APR-99.  
 DR PDB: 1E1L; 02-JUN-00.  
 DR InterPro: IPR000759; Adrndx\_reductse.  
 DR PRINTS; PK00419; AdrDXASE.  
 KW Electron transport; Oxidoreductase; Flavoprotein; NADP; FAD;  
 KW Mitochondrion; Transit peptide; Alternative splicing; 3D-structure.  
 FT TRANSIT 1 32 MITOCHONDRION.  
 FT CHAIN 33 492 NADPH:ADRENODOXIN OXIDOREDUCTASE.  
 FT VARSPIC 204 204 E -> EVLLCQ (IN LONG ISOFORM).  
 FT CONFLICT 77 77 G -> R (IN REF. 3).  
 FT CONFLICT 81 94 FGVAPDHPVKVNI -> VMLALTTPSRMLL (IN REF.  
 FT 3).  
 FT CONFLICT 124 128 QDAYH -> RYVRLT (IN REF. 3).  
 FT CONFLICT 268 268 PS -> R (IN REF. 3).  
 FT CONFLICT 317 318 K -> RL (IN REF. 3).  
 FT CONFLICT 323 333 RAAGIRLAVTR -> ARRSAQWSPQ (IN REF. 3).  
 FT CONFLICT 341 352 TRAVPTGVDVDEL -> HPSAHWGCGGP (IN REF. 3).  
 FT CONFLICT 341 352 TRAVPTGVDVDEL -> HPSAHWGCGGP (IN REF. 3).  
 SQ SEQUENCE 492 AA; 54338 MW; E68F6F5D18F53131 CRC64;  
 Query Match 50.0%; Score 51.5; DB 1; Length 492;  
 Best Local Similarity 63.6%; Pred. No. 5.5;  
 Matches 7; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 2 RRPWPPWP-KW 11  
 Db 4 RCWRWPPWSSW 14  
 RESULT 6  
 VGL2\_CVH22 VGL2\_CVH22 STANDARD; PRT; 1173 AA.  
 AC P15423;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN).  
 OS S.  
 GN Human coronavirus (strain 229E).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;  
 OC Coronaviridae; Coronavirus.  
 OX NCBI\_TaxID=111137;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90264837; PubMed=2345367;  
 RA Raabe T., Schelle-Prinz B., Siddell S.G.;  
 RT "Nucleotide sequence of the gene encoding the spike glycoprotein of  
 RT human coronavirus HCV 229E.";  
 RL J. Gen. Virol. 71:1065-1073(1990).  
 CC -!- FUNCTION: THE PELOMER PROTEIN MEDIATES THE BINDING OF VIRIONS  
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION  
 CC AND IN SYNCYTIIUM FORMATION.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X16816; CAA34723.1;  
 CC PIR: A34766; VGIHHC.  
 DR InterPro: IPR002551; Corona\_S1.  
 DR InterPro: IPR002552; Corona\_S2.  
 DR Pfam: PF01600; Corona\_S1; 1.  
 DR Pfam: PF01601; Corona\_S2; 1.  
 DR Glycoprotein; Envelope protein; Transmembrane; Signal.  
 KW Glycoprotein; 1 15  
 FT SIGNAL 16 1173  
 FT CHAIN 16 1115  
 FT DOMAIN 16 1115  
 FT TRANSMEM 1116 1135  
 FT POTENTIAL.



Thu Jan 31 11:07:41 2002

CC (SDR) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: Z79700; CAB02005.1; -

DR EMBL: AF006982; AAK45219.1; -

DR TIGR: MT0971; -

DR Tuberculin; RV0945; -

DR InterPro: IPR002198; ADH\_short.

DR Pfam: PF00106; adh\_short; 1.

DR PROSITE: PS00061; ADH\_SHORT; 1.

KW Hypothetical protein; Oxidoreductase; Complete proteome.

FT ACT\_SITE 159 159 BY SIMILARITY.

SQ SEQUENCE 253 AA; 27138 MW; BAD937208842DA12 CRC64;

Query Match 48.1%; Score 49.5; DB 1; Length 253;

Best Local Similarity 70.0%; Pred. No. 5.4; 1; Indels 1; Gaps 1;

Matches 7; Conservative 1; Mismatches 1;

QY 5 PWPWPWKWPLI 14

Db 230 PWPWPA-PLV 238

||||| ||:

RESULT 9

ID PF11\_PIG STANDARD; PRT; 212 AA.

AC P51524; 1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE PROPHENIN-1 PRECURSOR (PF-1) (C6) (FRAGMENT).

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=96042752; PubMed=7576250;

RA Strukelj B., Pungercar J., Kopitar G., Renko M., Lenarcic B.,

RA Berbic S., Turk V.;

RT "Molecular cloning and identification of a novel porcine

RL cathelin-like antibacterial peptide precursor.";

RL Biol. Chem. Hoppe-Seyler 376:507-510(1995).

RN [2]

RP SEQUENCE OF 131-209.

RC TISSUE=Blood;

RX MEDLINE=95212585; PubMed=7698355;

RA Harwig S.L., Kokryakov V.N., Swiderek K.M., Aleshina G.M.,

RA Zhao C., Lehrer R.I.;

RT "Prophenin-1, an exceptionally proline-rich antimicrobial peptide

RL from porcine leukocytes.";

RL FEBS Lett. 362:65-69(1995).

CC -!- FUNCTION: EXERTS ANTIMICROBIAL ACTIVITY. IT IS MORE EFFECTIVE

CC AGAINST GRAM NEGATIVE BACTERIA THAN GRAM POSITIVE BACTERIA.

CC -!- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X86031; CAA60023.1; -

DR InterPro: IPR001894; Cathelicidin.

DR Pfam: PF00666; Cathelicidins; 1.

DR ProDom: PD001838; Cathelicidin; 1.

DR PROSITE: PS00947; CATHELICIDINS\_2; 1.

DR PROSITE: PS00947; CATHELICIDINS\_2; 1.

KW Antibiotic; Repeat; Amidation; Signal.

FT NON\_TER 1 1

FT SIGNAL <1 13 POTENTIAL.

FT PROPEP 14 130 PROPHENIN-1.

FT PEPTIDE 131 209 REMOVED IN MATURE FORM (POTENTIAL).

FT PROPEP 210 212 PYRROLIDONE CARBOXYLIC ACID (BY

FT MOD\_RES 14 14 SIMILARITY).

FT DISULFID 69 80 BY SIMILARITY.

FT DISULFID 91 108 BY SIMILARITY.

FT MOD\_RES 209 209 AMIDATION (G-210 PROVIDE AMIDE GROUP)

FT DOMAIN 116 212 (POTENTIAL).

FT DOMAIN 132 201 PRO-RICH.

FT REPEAT 142 141 7 X 10 AA TANDEM REPEATS.

FT REPEAT 142 151 1.

FT REPEAT 152 161 2.

FT REPEAT 162 171 3.

FT REPEAT 162 171 4.

FT REPEAT 172 181 5.

FT REPEAT 172 181 6.

FT REPEAT 182 191 7.

FT REPEAT 192 201

SQ SEQUENCE 212 AA; 23956 MW; A315414C90DBF423 CRC64;

Query Match 47.1%; Score 48.5; DB 1; Length 212;

Best Local Similarity 53.8%; Pred. No. 6.1;

Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;

QY 2 RRWPWPWKWPLI 14

Db 113 RRPWW--WPFL 122

||||| ||:

RESULT 10

ID PF12\_PIG STANDARD; PRT; 228 AA.

AC P51525;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE PROPHENIN-2 PRECURSOR (PP-2) (C12) (PROPHENIN-1 LIKE).

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Bone marrow;

RX MEDLINE=94085623; PubMed=8262247;

RA Pungercar J., Strukelj B., Kopitar G., Renko M., Lenarcic B.,

RA Gubensek F., Turk V.;

RT "Molecular cloning of a putative homolog of proline/arginine-rich

RL antibacterial peptides from porcine bone marrow.";

RL FEBS Lett. 336:284-288(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=96105365; PubMed=7498526;

RA Zhao C., Ganz T., Lehrer R.I.;

RT "Structures of genes for two cathelin-associated antimicrobial

RL peptides: prophenin-2 and PR-39.";

RL FEBS Lett. 376:130-134(1995).

CC -!- FUNCTION: EXERTS ANTIMICROBIAL ACTIVITY. IT IS MORE EFFECTIVE

CC AGAINST GRAM NEGATIVE BACTERIA THAN GRAM POSITIVE BACTERIA.

CC -!- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X75438; CAA53188.1; -  
 EMBL; X89202; CAA61488.1; -  
 InterPro: IPR001894; Cathelicidin.  
 Pfam: PF00666; Cathelicidins; 1.  
 ProDom: PD001838; Cathelicidin; 1.  
 ProSITE; PS00946; CATHELICIDINS\_1; 1.  
 ProSITE; PS00947; CATHELICIDINS\_2; 1.  
 Antibiotic; Repeat; Amidation; Signal.  
 SIGNAL 1 29 POTENTIAL.  
 FT PROPEP 30 146 POTENTIAL.  
 FT PEPTIDE 147 225 PROPENIN-2.  
 FT PROPEP 226 228 REMOVED IN MATURE FORM (POTENTIAL).  
 FT MOD\_RES 225 225 AMIDATION (G-226 PROVIDE AMIDE GROUP) (POTENTIAL).  
 FT DOMAIN 132 228 PRO-RICH.  
 FT DOMAIN 148 217 7 X 10 AA TANDEM REPEATS.  
 FT REPEAT 148 157 1.  
 FT REPEAT 158 167 2.  
 FT REPEAT 168 177 3.  
 FT REPEAT 178 187 4.  
 FT REPEAT 188 197 5.  
 FT REPEAT 198 207 6.  
 FT REPEAT 208 217 7.  
 FT MOD\_RES 30 30 PYRROLIDONE CARBOXYLIC ACID (BY SIMILARITY).  
 FT DISULFID 85 96 BY SIMILARITY.  
 FT DISULFID 107 124 BY SIMILARITY.  
 SQ SEQUENCE 228 AA; 25855 MW; 1E44511FF35CC182 CRC64;

Query Match 47.1%; Score 48.5; DB 1; Length 228;  
 Best Local Similarity 53.8%; Pred. No. 6.6;  
 Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 1;  
 QY 2 RRPWPWPWKWPLI 14  
 DB 129 RRPWPWPWKWPLI 138

RESULT 11  
 ENV\_OMVVS  
 ID ENV\_OMVVS STANDARD; PRT; 990 AA.  
 AC P16899;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN).  
 GN ENV.  
 OS Ovine lentivirus (strain SA-OMVVS).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90223989; PubMed=2158181;  
 RA Querat G., Audoly G., Sonigo P., Vigne R.;  
 RT "Nucleotide sequence analysis of SA-OMVVS, a visna-related ovine lentivirus: phylogenetic history of lentiviruses";  
 RL Virology 175:434-447(1990).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M34193; AAA46783.1; -  
 DR EMBL; M31646; AAA66817.1; -  
 DR HIV; G46335; G46335.  
 DR PIR; M34193; ENV\$OMVVSACG.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam: PF00517; GP41; 1.  
 DR Glycoprotein; Coat protein; Polypeptide; Transmembrane.  
 KW Glycoprotein; Coat protein; Polypeptide; Transmembrane.  
 FT PEPTIDE 1 101 LEADER PEPTIDE.  
 FT CHAIN 102 662 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT TRANSMEM 663 990 TRANSMEMBRANE GLYCOPROTEIN.  
 FT PROSITE; PS00946; CATHELICIDINS\_1; 1.  
 FT PROSITE; PS00947; CATHELICIDINS\_2; 1.  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 490 490 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 703 703 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 771 771 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 778 778 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 794 794 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 990 AA; 114498 MW; 279B816B55614F3 CRC64;

Query Match 46.6%; Score 48; DB 1; Length 990;  
 Best Local Similarity 50.0%; Pred. No. 28;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 2 RRPWPWPWKWPLI 13  
 DB 174 QRPWPWPWKWPLI 185

RESULT 12  
 ENV\_CAEVG  
 ID ENV\_CAEVG STANDARD; PRT; 942 AA.  
 AC P31627;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: SURFACE PROTEIN; TRANSMEMBRANE PROTEIN].  
 GN ENV.  
 OS Caprine arthritis encephalitis virus (strain 663) (CAEV).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11662;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 81-95.  
 RX MEDLINE=92015464; PubMed=1656067;  
 RA Knowles D.P. Jr., Cheevers W.P., McGuire T.C., Brassfield A.L.,  
 RT "Structure and genetic variability of envelope glycoproteins of two antigenic variants of caprine arthritis-encephalitis lentivirus";  
 RL J. Virol. 65:5744-5750(1991).  
 RN [2]

REVISIONS.  
 Knowles D.P.;  
 Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M60855; AAB88709.2; -  
 PIR; A41307; VCLJCC6.  
 InterPro; IPR000328; Env\_GP41.  
 Pfam; PF00517; GP41; 1.  
 Glycoprotein; Coat protein; Polyprotein; Transmembrane.  
 LEADER PEPTIDE.  
 SURFACE PROTEIN (POTENTIAL).  
 TRANSMEMBRANE PROTEIN (POTENTIAL).  
 EXTRACELLULAR (POTENTIAL).  
 POTENTIAL.  
 CYTOPLASMIC (POTENTIAL).  
 TRANSMEM  
 DOMAIN 800 820  
 DOMAIN 821 942  
 CARBOHYD 51 51  
 CARBOHYD 98 98  
 CARBOHYD 131 131  
 CARBOHYD 176 176  
 CARBOHYD 228 228  
 CARBOHYD 331 331  
 CARBOHYD 348 348  
 CARBOHYD 354 354  
 CARBOHYD 370 370  
 CARBOHYD 379 379  
 CARBOHYD 400 400  
 CARBOHYD 404 404  
 CARBOHYD 435 435  
 CARBOHYD 441 441  
 CARBOHYD 447 447  
 CARBOHYD 457 457  
 CARBOHYD 467 467  
 CARBOHYD 481 481  
 CARBOHYD 493 493  
 CARBOHYD 503 503  
 CARBOHYD 509 509  
 CARBOHYD 527 527  
 CARBOHYD 534 534  
 SEQUENCE 942 AA; 108437 MW; 5B0E1F0F3D355F4A CRC64;  
 -----  
 Query Match 45.6%; Score 47; DB 1; Length 942;  
 Best Local Similarity 60.0%; Pred. No. 36;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 -----  
 QY 4 WPMWPKWPL 13  
 111 : 111  
 Db 145 WPMWPKWPL 154  
 -----  
 RESULT 13  
 ENV\_CAECV STANDARD; PRT; 966 AA.  
 AC P31626;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: SURFACE  
 DE PROTEIN; TRANSMEMBRANE PROTEIN].  
 GN ENV.  
 OS Caprine arthritis encephalitis virus (strain Cork) (CAEV).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11661;

SEQUENCE FROM N.A.  
 MEDLINE=92015464; PubMed=1656067;  
 Knowles D.P. Jr.; Cheevers W.P.; McGuire T.C.; Brassfield A.L.;  
 Harwood W.G.; Stem T.A.;  
 "Structure and genetic variability of envelope glycoproteins of two  
 antigenic variants of caprine arthritis-encephalitis lentivirus."  
 J. Virol. 65:5744-5750(1991).  
 -----  
 [2]  
 SEQUENCE FROM N.A.  
 MEDLINE=91021037; PubMed=2171210;  
 Saltarelli M.; Querat G.; Konings D.A.; Vigne R.; Clements J.E.;  
 "Nucleotide sequence and transcriptional analysis of molecular clones  
 of CAEV which generate infectious virus."  
 Virology 179:347-364(1990).  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 EMBL; M33677; AAA91829.1; -  
 PIR; B41307; VCLJCC.  
 InterPro; IPR000328; Env\_GP41.  
 Pfam; PF00517; GP41; 1.  
 Glycoprotein; Coat protein; Polyprotein; Transmembrane.  
 LEADER PEPTIDE.  
 SURFACE PROTEIN (POTENTIAL).  
 TRANSMEMBRANE PROTEIN (POTENTIAL).  
 EXTRACELLULAR (POTENTIAL).  
 POTENTIAL.  
 CYTOPLASMIC (POTENTIAL).  
 TRANSMEM  
 DOMAIN 662 801  
 DOMAIN 802 822  
 DOMAIN 823 966  
 CARBOHYD 101 101  
 CARBOHYD 134 134  
 CARBOHYD 179 179  
 CARBOHYD 220 220  
 CARBOHYD 231 231  
 CARBOHYD 334 334  
 CARBOHYD 351 351  
 CARBOHYD 357 357  
 CARBOHYD 382 382  
 CARBOHYD 403 403  
 CARBOHYD 407 407  
 CARBOHYD 438 438  
 CARBOHYD 443 443  
 CARBOHYD 449 449  
 CARBOHYD 459 459  
 CARBOHYD 469 469  
 CARBOHYD 483 483  
 CARBOHYD 495 495  
 CARBOHYD 505 505  
 CARBOHYD 511 511  
 CARBOHYD 529 529  
 CARBOHYD 536 536  
 CONFLICT 41 41  
 CONFLICT 47 47  
 CONFLICT 55 55  
 CONFLICT 834 834  
 CONFLICT 936 936  
 CONFLICT 941 966  
 SEQUENCE 966 AA; 110291 MW; 2ADEB6A92EC1D69B CRC64;  
 -----  
 Query Match 45.6%; Score 47; DB 1; Length 966;  
 Best Local Similarity 60.0%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 -----  
 QY 4 WPMWPKWPL 13  
 111 : 111  
 Db 145 WPMWPKWPL 154  
 -----  
 RESULT 13  
 ENV\_CAECV STANDARD; PRT; 966 AA.  
 AC P31626;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN) [CONTAINS: SURFACE  
 DE PROTEIN; TRANSMEMBRANE PROTEIN].  
 GN ENV.  
 OS Caprine arthritis encephalitis virus (strain Cork) (CAEV).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OC NCBI\_TaxID=11661;



QY 4 WPMWPKWPL 13  
 DB 148 WPMWPKWPL 157

RESULT 14  
 ENV\_VILV STANDARD; PRT; 982 AA.  
 ID ENV\_VILV STANDARD; PRT; 982 AA.  
 AC P03379;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DE 15-JUL-1999 (Rel. 36, Last annotation update)  
 EN ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN).  
 GN ENV.  
 OS Visna lentivirus (strain 1514).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11742;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85254938; PubMed=2410140;  
 RA Sonigo P., Allison M., Staskus K., Klatzmann D., Cole S., Danos O.,  
 RA Retzel E., Tiollais P., Haase A., Wain-Hobson S.;  
 RT "Nucleotide sequence of the visna lentivirus: relationship to the  
 RT AIDS virus.";  
 RL Cell 42:369-382(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88062965; PubMed=2824836;  
 RA Braun M.J., Clements J.E., Gonda M.A.;  
 RT "The visna virus genome: evidence for a hypervariable site in the env  
 RT gene and sequence homology among lentivirus envelope proteins.";  
 RL J. Virol. 61:4046-4054(1987).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M10608; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL; A15114; CAA01216.1; -;  
 DR PIR; A03977; VCLJVS.  
 DR HIV; M10608; ENVSVLVC.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 DR KW Glycoprotein; Coat protein; Polyprotein; Transmembrane.  
 FT PEPTIDE 1 100  
 FT CHAIN 101 656  
 FT CHAIN 657 982  
 FT TRANSMEM 835 851  
 FT CARBOHYD 140 140  
 FT CARBOHYD 161 161  
 FT CARBOHYD 206 206  
 FT CARBOHYD 258 258  
 FT CARBOHYD 298 298  
 FT CARBOHYD 364 364  
 FT CARBOHYD 370 370  
 FT CARBOHYD 381 381  
 FT CARBOHYD 387 387  
 FT CARBOHYD 403 403  
 FT CARBOHYD 414 414  
 FT CARBOHYD 435 435  
 FT CARBOHYD 439 439  
 FT CARBOHYD 470 470  
 FT CARBOHYD 475 475  
 FT CARBOHYD 481 481  
 FT CARBOHYD 491 491  
 FT CARBOHYD 501 501  
 FT CARBOHYD 515 515  
 FT CARBOHYD 527 527

FT CARBOHYD 537  
 FT CARBOHYD 542  
 FT CARBOHYD 543  
 FT CARBOHYD 568  
 FT CARBOHYD 597  
 FT CARBOHYD 764  
 FT CARBOHYD 771  
 FT CARBOHYD 787  
 FT CARBOHYD 821  
 FT VARIANT 12  
 FT VARIANT 118  
 FT VARIANT 283  
 FT VARIANT 640  
 FT VARIANT 645  
 SQ SEQUENCE 982 AA; 113978 MW; 7D78BA6E22BF53F CRC64;

Query Match 45.6%; Score 47; DB 1; Length 982;  
 Best Local Similarity 60.0%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 OY 4 WPMWPKWPL 13  
 DB 175 WPMWPKWPL 184

RESULT 15  
 ENV\_VILV STANDARD; PRT; 983 AA.  
 AC P35954;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DE ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN).  
 GN ENV.  
 OS Visna lentivirus (strain KVI772).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=36374;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93174981; PubMed=8382414;  
 RA Andreessen O.S., Elser J.E., Tobin G.J., Greenwood J.D., Gonda M.A.,  
 RA Georgsson G., Andresdottir V., Benediktsson E., Carlsson H.M.,  
 RA Maentylae E.O., Rafnar B., Palsson P.A., Casey J.W., Petursson G.;  
 RT "Nucleotide sequence and biological properties of a pathogenic  
 RT proviral molecular clone of neurovirulent visna virus.";  
 RL Virology 193:89-105(1993).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; S55323; AAB25463.1; -;  
 DR EMBL; L06906; AAM48362.1; -;  
 DR PIR; E45390; E45390.  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00517; GP41; 1.  
 DR KW Glycoprotein; Coat protein; Polyprotein; Transmembrane.  
 FT PEPTIDE 1 100  
 FT CHAIN 101 656  
 FT CHAIN 657 983  
 FT TRANSMEM 836 852  
 FT CARBOHYD 140 140  
 FT CARBOHYD 161 161  
 FT CARBOHYD 206 206  
 FT CARBOHYD 258 258  
 FT CARBOHYD 298 298  
 FT CARBOHYD 364 364  
 FT CARBOHYD 381 381

us-09-432-546-5.rsp

Thu Jan 31 11:07:41 2002

FT	CARBOHYD	387	387	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	403	403	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	414	414	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	435	435	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	439	439	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	470	470	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	475	475	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	481	481	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	491	491	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	501	501	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	515	515	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	527	527	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	537	537	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	542	542	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	543	543	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	568	568	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	697	697	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	765	765	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	772	772	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	788	788	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	822	822	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CONFLICT	900	900	R -> H (IN REF. 1: AAA48362).	
SQ	SEQUENCE	983 AA;	113939 MW;	E03BCC254EA78268	CRC64;

Query Match 45.6%; Score 47; DB 1; Length 983;  
 Best Local Similarity 60.0%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 WPMWPKWPL 13  
 III : III  
 Db 175 WPMWPKWPL 184

Search completed: January 30, 2002, 11:52:22  
 Job time: 201 sec